

The Cancer Genome Atlas Pilot Project

DATA USE CERTIFICATION

version date 02182009

Introduction and Statement of Policy

The National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) have launched The Cancer Genome Atlas (TCGA) Pilot Project to assess the feasibility of a full-scale effort to systematically explore the entire spectrum of genomic changes involved in human cancer. Implicit in TCGA is the view that scientific progress in this area will be greatly enhanced if the data that are produced by the program are readily available to the research community.

TCGA Datasets to be distributed under this certification process will be coded and distributed from the [TCGA Data Portal](#) or, for genome sequencing datasets, the National Center for Biotechnology Information Database of Genotype and Phenotype (dbGaP). However, due to the nature of sequence and genotype data, it may be possible, in principle, to use the data to identify an individual research participant. Thus, such data will only be provided to investigators who, along with their institutions, have certified their agreement with the requirements and terms of access and data use detailed below.

To promote the responsible use of TCGA Datasets, all investigators and their institutions seeking access to the data shall acknowledge their agreement with [TCGA Policies and Procedures](#) as described on the TCGA website, and which are articulated within this Data Use Certification.

Terms of Access

1. Requesting Investigator ("Requester")

The Requester has reviewed and understands the guiding principles for responsible research use and handling of the data included within TCGA Datasets as described in the on-line material provided through the [TCGA website](#) and [dbGaP](#).

2. Research Use

The Requester agrees that he/she shall use TCGA Datasets solely in connection with the research project described in the Data Access Request (DAR). New uses of these data outside those described in the DAR or substantive modifications to the research project will require submission of a new request. The Requester further agrees that he/she shall use TCGA Dataset(s) only in accordance with the parameters described on the [TCGA website](#) for appropriate research uses, and any limitations on such use, of this dataset. The Requester and his/her institution further acknowledge that they are responsible for ensuring that all uses of the data are consistent with federal, state, and local laws and regulations and any relevant institutional policies.

It is anticipated that TCGA Datasets will be updated with additional information. Unless otherwise indicated, all statements herein are presumed to be true and applicable to the access and use of all versions of TCGA Datasets.

3. Non-Identification

The Requester agrees not to use TCGA Datasets, or any other information, to identify or contact individual participants, or their relatives, from whom phenotype data and tissue samples were collected.

Exceptions: A Requester who is also a TCGA contributing investigator is exempt from this condition if explicit permission to integrate local identifying data has been obtained under a human subjects research protocol approved by the IRB of the contributing institution. Also, if the Requester is collaborating with a contributing investigator who has the potential to re-identify participants, IRB approval of the research protocol is required from the contributing institution. Implicit in submission of the DAR, the institutional signing official provides assurance that IRB approval has been obtained in such a circumstance.

4. Non-transferability

The Requester agrees to retain control over the data, and further agrees not to distribute the individual data in any form to any entity or individual other than research assistants, students, post-doctoral fellows, or other research staff whose work is supervised directly by the Requester and who agree in writing to the terms of this Data Use Certification. The Requester and his/her institution accept responsibility for ensuring appropriate use of these data and obtaining agreement to the terms of this document by research staff.

The Requester agrees that if he/she changes institutions, then a new Data Access Request incorporating agreement to the Data Use Certification by the new institution will be submitted and approved before data use resumes. Any versions of data stored at the prior institution will be destroyed.

5. Computer Security

The Requester agrees to store TCGA data only on computers with adequate security controls, to adhere to practices that ensure that only authorized individuals can gain access, and to maintain appropriate control over any copies of the data.

In the event of a breach of confidentiality, Approved Users are required to report the breach on the same day it is discovered. Reports can be made by contacting the NCI Center for Bioinformatics Application Support by phone (toll free 888-478-4423 or 301-451-4384) or email (ncicb@pop.nci.nih.gov). The report must include all known information regarding the breach as well as a description of plans for further investigation and immediate remediation.

The NCI, NHGRI or other entity designated by the NCI or NHGRI may also investigate such breaches. Approved Users and their associates will be required to support such investigations and provide information, within the limits of applicable local, state and federal laws and regulations.

6. Intellectual Property

By requesting access to TCGA Dataset(s), the Requester and his/her Institution acknowledge the intent of NCI and NHGRI to see that Approved Users follow the TCGA intellectual property policy as summarized below:

- Achieving maximum public benefit is the ultimate goal of TCGA. NCI/NHGRI believe that TCGA data should be considered as pre-competitive, and urge users to avoid making claims to intellectual property (IP) derived directly from the data. However, NCI/NHGRI also recognize the importance of the later development of IP on downstream discoveries, especially in therapeutics, which will be necessary to support full investment in and development of products that the public needs.
- It is expected that TCGA-supported data and conclusions derived therefrom will remain freely available, without requirement for licensing. The NCI and NHGRI encourage broad use of TCGA data coupled with a responsible approach to management of IP derived from downstream discoveries that are consistent both with the recommendations cited in NIH's [Best Practices for the Licensing of Genomic Inventions](#) and in the [NIH Research Tools Policy](#).

7. Publication

It is the intent of NCI and NHGRI to promote the dissemination of analyses of TCGA Dataset(s) as widely as possible. Approved Users are strongly encouraged to publish their results in peer-reviewed journals, but are asked to apply the normal standards of scientific etiquette when deciding to publish results based substantially on unpublished TCGA data. NCI and NHGRI intend that TCGA genomic data are released as rapidly as possible after they are produced, with no restrictions on use. However, Approved Users will agree that any publication of results or analyses derived from the use of TCGA Datasets will acknowledge TCGA with the following sentence:

“The results published here are in whole or part based upon data generated by The Cancer Genome Atlas pilot project established by the NCI and NHGRI. Information about TCGA and the investigators and institutions who constitute the TCGA research network can be found at [“http://cancergenome.nih.gov”](http://cancergenome.nih.gov).

8. Research Use Reporting

To ensure that TCGA policies and procedures for participant protection and data use are adhered to, Approved Users agree to provide annual feedback on how they have used TCGA data and any results such as publications and patents that have been generated as a result of access to the data. This information will be used by the TCGA Data Use Review Board (DURB) as part of its effort to provide ongoing, independent oversight of all TCGA Dataset use and by NCI/NHGRI staff as required for program oversight and management activities.

Approved Users thus agree to provide a brief report on the research specified within this DAR. Requesters who were previously Approved Users (that is, those seeking renewal) will be asked to provide specific information in the renewal DAR. Those not seeking renewal will be asked to complete only a specific reporting section of the DAR. The report will be in an automated format, and will ask questions regarding publications that resulted from the use of TCGA data, breaches in security of data (for example, accidental distribution of data beyond Approved Users), information on any intellectual property generated as a result of the data, and any return of research results to research participants (in cases where the Approved User is also a Contributing Investigator). Approved Users may be asked to provide general comments regarding the effectiveness of TCGA, such as ease of access and use, appropriateness of data format, challenges in complying with TCGA policies, and suggestions for improving TCGA data access or the program in general.

9. Non-Endorsement, Indemnification

The Requester and his/her institution acknowledge that although all reasonable efforts have been taken to ensure the accuracy and reliability of TCGA data, NIH, NCI, NHGRI, all Centers in the TCGA research network, and Contributing Investigators do not and cannot warrant the results that may be obtained by using any data included therein. NIH, NCI, NHGRI and all contributors to the TCGA datasets disclaim all warranties as to performance or fitness of the data for any particular purpose.

No indemnification for any loss, claim, damage or liability is intended or provided by any party under this agreement. Each party shall be liable for any loss, claim, damage, or liability that said party incurs as a result of its activities under this agreement, except that the NIH, as an agency of the United States, assumes liability only to the extent provided under the Federal Tort Claims Act, 28 U.S.C. 2671 et seq.

10. Termination and Violations

This Data Use Certification will be in effect for a period of one year from the date the data are made accessible to the Approved User. At the end of this time period, the Approved User agrees to destroy all copies of TCGA controlled-access datasets and all derivatives that contain individual-level (potentially identifying) information unless access rights are renewed upon submission of a new DAR.

NCI/NHGRI may terminate this agreement and immediately revoke access to all TCGA controlled-access Datasets at any time if the Requester is found to be no longer in compliance with the policies, principles and procedures of TCGA.

By submission of a Data Access Request form, the Requester and Authorized Institutional Official represent and warrant the Requester's qualifications for access to and use of TCGA Datasets and certify their agreement to the TCGA and dbGaP principles, policies and

procedures for the use of TCGA Datasets as articulated in this document. Requesters further acknowledge that they have shared this document and TCGA policies and procedures with any research staff who will participate in the use of the TCGA Dataset. Authorized Institutional Officials also acknowledge that they have considered the relevant TCGA policies and procedures and have assured compliance with local institutional policies related to technology transfer and human subject's research.